

# Predição de Séries Temporais Baseada em Redes Neurais Artificiais

Aluno: João Pedro de Oliveira Pagnan [FEEC/UNICAMP]

Orientador: Prof. Dr. Levy Boccato [FEEC/UNICAMP]

Coorientador: Prof. Dr. Romis Ribeiro de Faissol Attux [FEEC/UNICAMP]

Departamento de Engenharia de Computação e Automação Industrial – DCA

26 de outubro de 2021

# Introdução

## Previsão de séries temporais de sistemas caóticos

O desafio de antecipar padrões de comportamento e construir modelos que sejam apropriados para explicar determinados fenômenos da natureza tem importância para várias áreas da ciência.

# Introdução

## Previsão de séries temporais de sistemas caóticos

Uma classe de sistemas dinâmicos particularmente relevante dentro do contexto de modelagem e predição de séries temporais está ligada à ideia de dinâmica caótica. Apesar de serem determinísticos, esses sistemas são extremamente sensíveis às condições iniciais<sup>[1]</sup>.

# Introdução

## Previsão de séries temporais com redes neurais artificiais

Tendo em vista o desempenho de modelos não-lineares para previsão de diversas séries temporais<sup>[2]</sup>, optamos por estudar a aplicabilidade de redes neurais artificiais à previsão de séries relacionadas a sistemas com dinâmica caótica.

Esta pesquisa comparou o desempenho de quatro arquiteturas de redes neurais artificiais:

- *Multilayer Perceptron (MLP)*<sup>[3]</sup>,

# Introdução

## Previsão de séries temporais com redes neurais artificiais

Tendo em vista o desempenho de modelos não-lineares para previsão de diversas séries temporais<sup>[2]</sup>, optamos por estudar a aplicabilidade de redes neurais artificiais à previsão de séries relacionadas a sistemas com dinâmica caótica.

Esta pesquisa comparou o desempenho de quatro arquiteturas de redes neurais artificiais:

- *Multilayer Perceptron (MLP)*<sup>[3]</sup>,
- *Long Short-Term Memory (LSTM)*<sup>[2]</sup>,

# Introdução

## Previsão de séries temporais com redes neurais artificiais

Tendo em vista o desempenho de modelos não-lineares para previsão de diversas séries temporais<sup>[2]</sup>, optamos por estudar a aplicabilidade de redes neurais artificiais à previsão de séries relacionadas a sistemas com dinâmica caótica.

Esta pesquisa comparou o desempenho de quatro arquiteturas de redes neurais artificiais:

- *Multilayer Perceptron (MLP)*<sup>[3]</sup>,
- *Long Short-Term Memory (LSTM)*<sup>[2]</sup>,
- *Gated Recurrent Unit (GRU)*<sup>[4]</sup>,

# Introdução

## Previsão de séries temporais com redes neurais artificiais

Tendo em vista o desempenho de modelos não-lineares para previsão de diversas séries temporais<sup>[2]</sup>, optamos por estudar a aplicabilidade de redes neurais artificiais à previsão de séries relacionadas a sistemas com dinâmica caótica.

Esta pesquisa comparou o desempenho de quatro arquiteturas de redes neurais artificiais:

- *Multilayer Perceptron (MLP)*<sup>[3]</sup>,
- *Long Short-Term Memory (LSTM)*<sup>[2]</sup>,
- *Gated Recurrent Unit (GRU)*<sup>[4]</sup>,
- *Echo State Network (ESN)*<sup>[5]</sup>.

# Introdução

## Cenários escolhidos

A comparação foi realizada considerando quatro cenários, sendo dois destes a tempo discreto e dois a tempo contínuo, sendo que foi gerado um conjunto de 5000 amostras para cada sistema analisado. Os cenários escolhidos foram as séries temporais dos sistemas caóticos:

- Mapa de Hénon<sup>[6]</sup>,

# Introdução

## Cenários escolhidos

A comparação foi realizada considerando quatro cenários, sendo dois destes a tempo discreto e dois a tempo contínuo, sendo que foi gerado um conjunto de 5000 amostras para cada sistema analisado. Os cenários escolhidos foram as séries temporais dos sistemas caóticos:

- Mapa de Hénon<sup>[6]</sup>,
- Mapa logístico<sup>[7]</sup>,

# Introdução

## Cenários escolhidos

A comparação foi realizada considerando quatro cenários, sendo dois destes a tempo discreto e dois a tempo contínuo, sendo que foi gerado um conjunto de 5000 amostras para cada sistema analisado. Os cenários escolhidos foram as séries temporais dos sistemas caóticos:

- Mapa de Hénon<sup>[6]</sup>,
- Mapa logístico<sup>[7]</sup>,
- Sistema de Lorenz<sup>[8]</sup>,

# Introdução

## Cenários escolhidos

A comparação foi realizada considerando quatro cenários, sendo dois destes a tempo discreto e dois a tempo contínuo, sendo que foi gerado um conjunto de 5000 amostras para cada sistema analisado. Os cenários escolhidos foram as séries temporais dos sistemas caóticos:

- Mapa de Hénon<sup>[6]</sup>,
- Mapa logístico<sup>[7]</sup>,
- Sistema de Lorenz<sup>[8]</sup>,
- Equações de Mackey-Glass<sup>[9]</sup>.

# Redes neurais utilizadas para a previsão

## MLP

As redes MLP, como o nome já diz, são compostas por múltiplas camadas de neurônios artificiais do tipo *perceptron*<sup>[3]</sup>. Cada neurônio recebe os atributos  $x_i, i = 1, \dots, m$  de entrada e os pondera por pesos  $w_i$ . Em seguida, é aplicada uma função de ativação  $\varphi(\cdot)$  sobre esta soma ponderada, acrescida de um termo de *bias*, conforme indica a equação (1).

$$y = \varphi \left( \sum_{i=1}^m w_i x_i + w_0 \right) = \varphi \left( \sum_{i=0}^m w_i x_i \right) = \varphi (w^T \cdot x) \quad (1)$$

# Redes neurais utilizadas para a previsão

## MLP

Uma rede neural MLP é composta por um número arbitrário  $N_L$  de camadas com  $n$  neurônios do tipo *perceptron*, com a característica de que as saídas dos neurônios da  $l$ -ésima camada são propagadas para a frente, servindo como as entradas de todos os neurônios da camada seguinte ( $l + 1$ ). A figura 1 apresenta a estrutura típica das redes MLP.

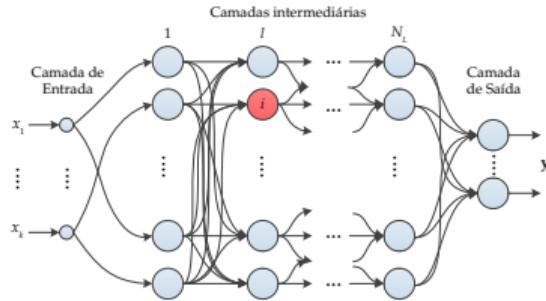


Figura: Estrutura típica de uma rede MLP (figura adaptada de<sup>[10]</sup>)

# Redes neurais utilizadas para a previsão

## MLP

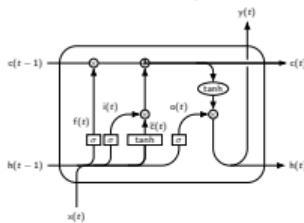
Os pesos sinápticos  $w$  são ajustados com um processo iterativo de forma a minimizar uma função custo  $J(w)$  que representa uma medida do erro entre as saídas geradas pela rede e as saídas desejadas.

# Redes neurais utilizadas para a previsão

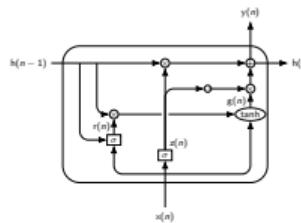
## LSTM e GRU

As redes recorrentes têm estruturas computacionais que podem armazenar os estados anteriores dos neurônios, possuindo também portas não-lineares que regulam o fluxo de informação de entrada e de saída da célula computacional [12].

A diferença entre uma célula LSTM e uma célula GRU é a presença de um vetor de longo prazo  $c$  e um vetor de curto prazo  $h$  na LSTM, conforme indicado na figura 2.



**Figura:** Estrutura de uma célula LSTM.



**Figura:** Estrutura de uma célula GRU.

# Redes neurais utilizadas para a previsão

## LSTM e GRU

À semelhança das redes MLP, o treinamento das redes LSTM e GRU também é realizado através de algoritmos de otimização baseados em derivadas da função custo; a diferença é que agora é necessário propagar as derivadas ao longo da estrutura e, também, ao longo do tempo devido às realimentações<sup>[11]</sup>.

## Redes neurais utilizadas para a previsão ESN

As redes neurais com estados de eco também são modelos recorrentes para processamento da informação, à semelhança da LSTM e da GRU. No entanto, apresentam um modo de operação e um esquema de treinamento bem diferentes [10].

A figura 4 apresenta a estrutura interna de uma rede ESN.

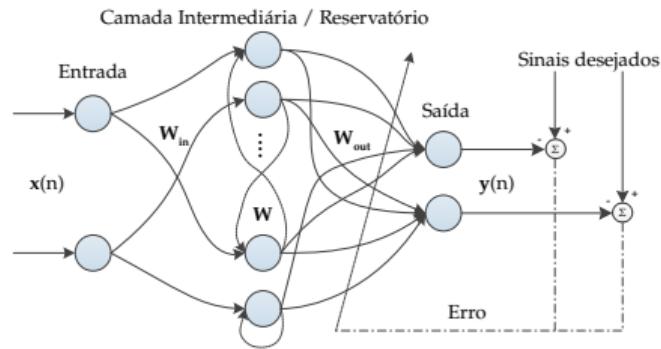


Figura: Estrutura típica de uma rede ESN (figura adaptada de<sup>[10]</sup>).

# Redes neurais utilizadas para a previsão ESN

O grande diferencial prático desta rede é que os pesos das conexões recorrentes dentro do reservatório  $W$  são ajustados com valores fixos antes do treinamento da camada de saída. Esses parâmetros são determinados tendo em vista a propriedade de estados de eco, existindo alguns métodos simples para a criação aleatória dos pesos e que asseguram a propriedade, conforme demonstrado em<sup>[5]</sup>.

# Metodologia

Iniciamos a análise através de um processo de busca em grade para determinarmos os parâmetros ótimos das redes neurais em cada cenário.

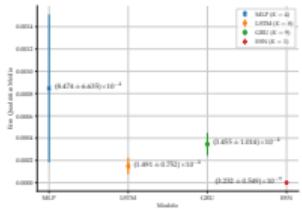
Em seguida, avaliamos a progressão do erro quadrático médio (EQM) em função do número de amostras de entrada do modelo preditor (chamado de  $K$ ).

# Metodologia

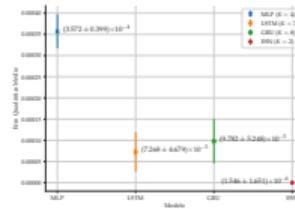
Por fim, comparamos qual foi a média e o desvio padrão do EQM com o melhor valor de  $K$  de cada modelo nos quatro cenários. O horizonte de predição utilizado foi  $L = 3$ . Assim, iremos predizer o valor da terceira iteração à frente do valor atual da série temporal.

# Resultados

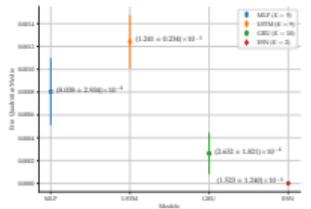
A figura 5 exibe o comparativo dos desempenhos obtidos por cada modelo com as melhores configurações nos cenários testados.



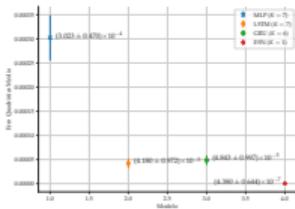
(a) Mapa de Hénon



(b) Mapa logístico



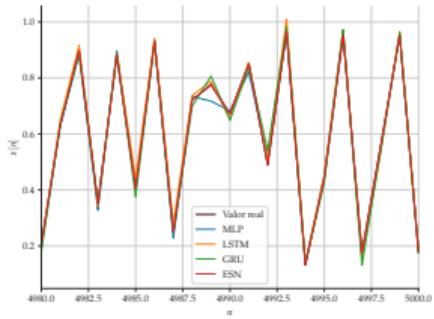
(c) Sistema de Lorenz



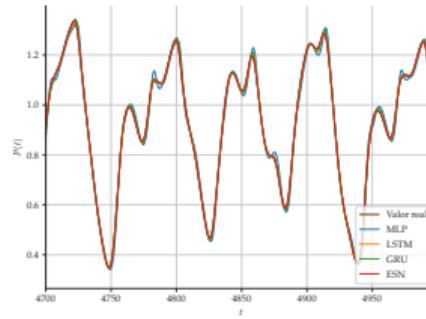
(d) Equações de Mackey-Glass

## Resultados

A figura 6 mostra uma comparação da predição em si de cada modelo nos cenários onde a diferença foi mais perceptível visualmente.



(a) Mapa logístico



(b) Equações de Mackey-Glass

**Figura:** Comparação da predição realizada por cada modelo nos cenários do mapa logístico e das equações de Mackey-Glass.

## Análise e Conclusão

Analisando os resultados obtidos, tiramos algumas conclusões importantes:

- A rede MLP foi consideravelmente pior do que as redes recorrentes,

## Análise e Conclusão

Analisando os resultados obtidos, tiramos algumas conclusões importantes:

- A rede MLP foi consideravelmente pior do que as redes recorrentes,
- Dentre as redes recorrentes, a ESN obteve um EQM bem inferior ao obtido pela LSTM e pela GRU,

## Análise e Conclusão

Analisando os resultados obtidos, tiramos algumas conclusões importantes:

- A rede MLP foi consideravelmente pior do que as redes recorrentes,
- Dentre as redes recorrentes, a ESN obteve um EQM bem inferior ao obtido pela LSTM e pela GRU,
- **A ESN atingiu patamares de erro cerca de 100 ou até 10000 vezes menores que os outros modelos.**

## Análise e Conclusão

O desempenho inferior da rede MLP com relação às redes recorrentes provavelmente decorre do fato de que a relação temporal presente na LSTM, GRU e ESN auxilia na modelagem da dinâmica da série temporal. Já o pior desempenho da LSTM na série temporal do sistema de Lorenz provavelmente está relacionado aos efeitos estocásticos presentes no ajuste dos pesos sinápticos dessa rede neural que, conforme indicado em [13], é uma dificuldade em seu treinamento.

## Análise e Conclusão

Considerando que a ESN obteve, de longe, o melhor desempenho, e que o treinamento desta arquitetura é bem menos custoso computacionalmente se comparado às demais, pode-se concluir que esta rede é uma boa alternativa para futuros estudos de modelos preditores de séries temporais.

## Análise e Conclusão

Como sugestão de trabalhos futuros, pode ser avaliada a eficácia da ESN em reconstruir atratores através de séries temporais experimentais de sistemas caóticos. Se o desempenho para essa tarefa for tão bom quanto o obtido na predição das séries estudadas, a ESN pode tornar-se uma ferramenta ainda mais poderosa para a modelagem de sistemas com dinâmica caótica.

# Referências

- [1] N. Fiedler-Ferrara and C. P. C. do Prado, *Caos: uma introdução*. Editora Blucher, 1994.
- [2] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 240–254, 1994.
- [3] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.,," *Psychological review*, vol. 65, no. 6, p. 386, 1958.
- [4] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [5] H. Jaeger, "Echo state network," *scholarpedia*, vol. 2, no. 9, p. 2330, 2007.
- [6] M. Hénon, "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, vol. 50, pp. 69–77, feb 1976.
- [7] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, pp. 459–467, jun 1976.
- [8] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [9] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, no. 4300, pp. 287–289, 1977.
- [10] L. Boccato, *Novas propostas e aplicações de redes neurais com estados de eco*. Tese (doutorado), Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP, 2013.
- [11] S. Haykin, *Neural networks and learning machines*, 3/E. Pearson Education India, 2010.
- [12] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.
- [13] K. Doya et al., "Bifurcations in the learning of recurrent neural networks 3," *learning (RTRL)*, vol. 3, p. 17, 1992.

Muito Obrigado!